

# Cloning, sequencing, and expression of the adenosylcobalamin-dependent ribonucleotide reductase from *Lactobacillus leichmannii*

(DNA biosynthesis/protein radicals)

SQUIRE BOOKER AND JOANNE STUBBE

Departments of Chemistry and Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

Contributed by JoAnne Stubbe, May 6, 1993

**ABSTRACT** Ribonucleoside-triphosphate reductase (RTPR, EC 1.17.4.2) from *Lactobacillus leichmannii*, a monomeric adenosylcobalamin-requiring enzyme, catalyzes the conversion of nucleoside triphosphates to deoxynucleoside triphosphates. The gene for this enzyme has been cloned and sequenced. In contrast to expectations based on mechanistic considerations, there is no statistically significant sequence homology with the *Escherichia coli* reductase that requires a dinuclear-iron center and tyrosyl radical cofactor. The RTPR has been overexpressed and purified to homogeneity, yielding 90 mg of protein from 2.5 g of bacteria. Initial characterization of the recombinant RTPR indicates that its properties are identical to those of the RTPR isolated from *L. leichmannii*.

Ribonucleotide reductases are uniquely responsible for converting nucleotides to deoxynucleotides *in vivo* (1–4). In contrast to most enzymes that play essential roles in metabolism in both prokaryotes and eukaryotes, the reductases do not appear, at least superficially, to have been evolutionarily conserved. In *Escherichia coli*, mammals, and herpesviruses, the enzymes possess two subunits, R1 and R2, each of which is homodimeric ( $\alpha_2\beta_2$ ). The R2 subunit possesses a dinuclear iron center-tyrosyl radical cofactor that is essential for nucleotide reduction. The R1 subunit contains the binding sites for the NDP substrates and dNTP allosteric effectors, and the cysteines that are oxidized concomitant with substrate reduction. The ribonucleoside-triphosphate reductase (RTPR, EC 1.17.4.2) from *Lactobacillus leichmannii*, the structurally simplest of all known reductases, is a monomer and requires adenosylcobalamin (AdoCbl) as a cofactor (5).

Despite these diverse cofactors and quaternary structures, evidence suggests that the *E. coli* and *L. leichmannii* enzymes exhibit common patterns of regulation and chemical mechanisms of reduction. The similarities in the allosteric regulatory properties are evident from investigations of the requirement for specific dNTPs to affect the reduction of specific purine and pyrimidine nucleotides (2, 6, 7). The similarities in mechanism of reduction are evident from investigations employing isotopically labeled nucleotide substrates and from investigations of the mechanisms of inhibition by 2'-chloro-2'-deoxynucleotides (3). With respect to the mechanism of reduction, both enzymes are proposed to initiate catalysis by generation of a protein radical that abstracts the hydrogen atom from the 3' carbon of the nucleotide substrate (3). The difference between the two enzymes is most apparent when considering the manner in which the two distinctly different cofactors give rise to the protein radical. For the *E. coli* reductase, it is proposed that the tyrosyl radical on the R2 subunit generates by long-range electron transfer a protein radical on the R1 subunit, which initiates turnover of the nucleotide (1, 8). The *L. leichmannii*

reductase is proposed to use AdoCbl as the equivalent of the R2-tyrosine radical. Homolysis of the carbon-cobalt bond is proposed to generate cob(II)alamin and a putative 5'-deoxyadenosyl radical. The latter is proposed to generate a protein radical that initiates turnover (9). To date, however, the dinuclear iron center-tyrosyl radical reductases, with *E. coli* as the prototype, have provided the major focus of interest for structural and molecular biological studies. Sequences of seven reductases (both the R1 and R2 subunits) in this family have now been reported (6). The structure of the R2 subunit of the *E. coli* ribonucleoside-diphosphate reductase (RDPR, EC 1.17.4.1) has been determined by crystallographic methods (8). Site-directed mutagenesis studies of conserved cysteines on the R1 subunit have begun to define residues that play an important role in catalysis (10–14). The unavailability of the genes has precluded these types of studies on the other classes of reductases.

The similarities in the chemistry and the allosteric regulation between the *E. coli* and *L. leichmannii* reductases, in spite of the differences in their quaternary structures and cofactor requirements, have motivated us to clone and sequence the gene for the *L. leichmannii* enzyme.\* The *L. leichmannii* reductase has also been overexpressed and purified to homogeneity, providing the foundation for further structural and mechanistic characterization.

## MATERIALS AND METHODS

Immobilon [poly(vinylidene difluoride)] membranes (0.45- $\mu$ m pore) were purchased from Millipore. Nitrocellulose membranes (BA-85, 0.45- $\mu$ m pore) were purchased from Schleicher & Schuell. Polyethylene glycol (PEG) 20M was from Baxter Scientific Products (McGaw Park, IL). Lysozyme (specific activity, 63,800 units/mg), ampicillin, *N*-laurylsarcosine, mutanolysin (specific activity, 6600 units/mg), ethidium bromide, Ponceau S, protamine sulfate, ATP, and AdoCbl were from Sigma. [ $\gamma$ - $^{32}$ P]ATP (6000 Ci/mmol; 1 Ci = 37 GBq) was purchased from New England Nuclear. Nick columns were from Pharmacia. Dithiothreitol was purchased from Mallinckrodt. Bio-Gel HTP hydroxylapatite was from Bio-Rad. DE52 anion-exchange resin was from Whatman. Kanamycin and Sequenase were purchased from United States Biochemical. The dsDNA Cycle Sequencing system, T4 DNA ligase, and competent DH5 $\alpha$  and HB101 cells were from GIBCO/BRL. AmpliTaq DNA polymerase was from Perkin-Elmer/Cetus. The TA cloning system was purchased from Invitrogen. Centricons, ultrafiltration devices, and membranes were purchased from Amicon. Restriction endonucleases and T4 DNA ligase were purchased from New England Biolabs. *E. coli* thioredoxin (specific activity, 50 units/mg) and thioredoxin

Abbreviations: RTPR, ribonucleoside-triphosphate reductase; AdoCbl, adenosylcobalamin; RDPR, ribonucleoside-diphosphate reductase.

\*The sequence reported in this paper has been deposited in the GenBank data base (accession no. L20047).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

reductase (specific activity, 800 units/mg) were isolated from overproducing strains SK3981 (15) and K91/pMR14 (16). Chromosomal DNA was isolated from *L. leichmannii* (ATCC 7830) grown in *Lactobacillus* carrying medium (17) supplemented with 1% dextrose and 20 mM D,L-threonine.

Peptide mapping and sequencing were done at the Harvard University Microchemistry Facility. Oligonucleotide primers used for DNA sequencing and PCR were obtained from the MIT Biopolymers Laboratory, or Oligos Etc. of Wilsonville, OR.

**Peptide Mapping.** RTPR (0.5 mg; specific activity, 1.3 units/mg) was subjected to SDS/PAGE in a 3-mm-thick 10% gel matrix. The gel was removed and soaked in transfer buffer [10 mM 3-(cyclohexylamino)-1-propanesulfonic acid, pH 11/10% methanol] for 30 min, assembled in a Hoeffer TE 70 SemiPhor semi-dry transfer unit, and blotted to a nitrocellulose membrane at 100 mA (constant current) for 3 h. The nitrocellulose membrane was removed and stained for 90 s in 0.2% Ponceau S/1% acetic acid, and destained for 90 s in 1% acetic acid. The protein band corresponding to RTPR was excised with a razor blade, rinsed with water, and submitted for peptide mapping (18, 19).

**Isolation of Genomic DNA.** Genomic DNA from *L. leichmannii* was purified by a modification of the procedure of Chassy and Guiffreda (20). One liter of *L. leichmannii* was grown at 37°C without aeration to an OD<sub>600</sub> of  $\approx 7$ . The cells were pelleted by centrifugation, washed in 0.02 M Tris-HCl (pH 8.2), and then resuspended in 25 ml of the same buffer. Lysozyme (730 mg in 20 ml of the above buffer), PEG [ $M_r$ , 20,000; 50 ml of a 24% (wt/vol) solution], and mutanolysin (0.5 mg in 0.5 ml of the above buffer) were added to the resuspended cells, and the resulting mixture was incubated for 1.5 h at 37°C. The mixture was then centrifuged at 5000  $\times g$  for 20 min. The supernatant was discarded and the pellet was resuspended in 40 ml of TE (10 mM Tris-HCl, pH 8.0/1 mM EDTA). A 5-ml aliquot of 10% (wt/vol) N-laurylsarcosine was added, and the resulting solution was incubated with agitation at 65°C for 15 min. Forty grams of CsCl and 4 ml of ethidium bromide at 10 mg/ml were added to the supernatant, and the DNA was purified using ultracentrifugation. The genomic DNA was isolated (21) and ultimately redissolved in TE at 1.2 mg/ml.

**Polymerase Chain Reaction (PCR).** Degenerate primers for the PCR were designed from regions of known amino acid sequence (Table 1) and were engineered with restriction sites at their 5' ends to facilitate the cloning of the amplified fragment. Primer 1 was generated from the N-terminal sequence of RTPR (AEFIDR), and primer 2 was generated from the putative active site region of RTPR (DQTDCE) (22). Primer 1, 5'-CGC-GGA-TCC-GC(G/T/A)-GA(G/A)-TT(C/T)-AT(C/T/A)-GA(C/T)-(C/A)G-3'; primer 2, 5'-GCG-GAA-TTC-(G/A)CA-(G/A)TC-(C/G/T/A)GT-(C/T)TG-(G/A)TC-3'. The PCR mixture contained (in a total volume of 100  $\mu$ l) 0.5–1.5  $\mu$ g of genomic DNA, 58 pmol of primer 1, 29 pmol of primer 2, all four dNTP (each at 0.2 mM), and 10  $\mu$ l of 10 $\times$  PCR buffer (500 mM KCl/100 mM Tris-HCl pH 8.3/15 mM MgCl<sub>2</sub>/0.1% gelatin). The mixture was overlaid with 100  $\mu$ l of paraffin oil and heated at 94°C for 5 min.

*Taq* polymerase (2.5 units) was added under the oil layer, and 35 cycles of the following program were run: 1 min at 94°C, 30 s at 37°C, 15 s at 50°C, and 2 min at 72°C. Upon completion, the paraffin oil was extracted with 150  $\mu$ l of CHCl<sub>3</sub>, and the DNA was precipitated. The DNA was dissolved in 440  $\mu$ l of H<sub>2</sub>O and desalted by several dilutions and centrifugations in a Centricon 100. To complete any unfinished sequences generated during the amplification, the PCR mixture was treated with Sequenase in the following manner. The reaction contained (in a final volume of 500  $\mu$ l) all four dNTPs (each at 0.1 mM), 1 $\times$  *Hind*III restriction buffer (50 mM NaCl/10 mM Tris-HCl, pH 7.9/10 mM MgCl<sub>2</sub>/1 mM dithiothreitol), 10 units of Sequenase, and the PCR product. The reaction was incubated at 30°C for 30 min and stopped by the addition of EDTA to a final concentration of 10 mM. The DNA was purified by electrophoresis in a 1% agarose gel, and the 2.1-kb piece was isolated by electroelution. The DNA was concentrated and exchanged into TE by using a Centricon 100, and its final concentration was determined using the relationship that 1 A<sub>260</sub> unit corresponds to 50  $\mu$ g/ml.

**Cloning and Sequencing of the Fragment Isolated by PCR.** The blunt-ended PCR fragment was ligated into pUC19 that had been digested with *Sma* I. The reaction mixture contained (in a final volume of 10  $\mu$ l) 100 ng of pUC19, 288 ng of the blunt-ended PCR fragment, 1  $\mu$ l of 10 $\times$  ligase buffer [500 mM Tris-HCl, pH 7.8/100 mM MgCl<sub>2</sub>/200 mM dithiothreitol/10 mM ATP/bovine serum albumin (500  $\mu$ g/ml)], and 7 Weiss units of T4 DNA ligase. The reaction mixture was incubated overnight at 16°C and then diluted to 200  $\mu$ l with TE buffer. A 5- $\mu$ l aliquot was used to transform competent *E. coli* DH5 $\alpha$ . The transformation reaction was plated on SOC agar (21) containing 0.008% 5-bromo-4-chloro-3-indolyl  $\beta$ -D-galactoside, and recombinants were identified by their white phenotype. Plasmid DNA was isolated from overnight cultures of several of the white colonies and screened for inserts by standard methods.

DNA sequencing was carried out using the dsDNA Cycle Sequencing system. The universal primer was used to initiate the sequencing process. All subsequent primers (21–24 bases long) were designed using the new sequence data from each successive round of sequencing. Approximately 200–300 ng of either the cloned or uncloned PCR fragment and 1.5–5  $\mu$ g of genomic DNA were used for each sequencing reaction mixture. The sequence of both strands of the uncloned PCR fragment was determined.

**Expression of RTPR.** Peptide sequence information from the N terminus and C terminus of RTPR was used to generate the appropriate PCR primers. The N-terminal primer 3, which was 49 bases long, contained an *Eco*RI restriction site, a consensus ribosome binding site, an 8-base transitional spacing element, and the first 24 bases of the RTPR gene sequence. The C-terminal primer 4, which was 31 bases long, contained a *Hind*III restriction site and the last 21 bases of the RTPR gene sequence, including the stop codon. Sequences of primers 3 and 4 are as follows: Primer 3, 5'-GCC-GGC-GAA-TTC-AGG-AGA-AAA-TAT-TAT-GAG-TGA-AGA-AAT-ATC-TCT-CTC-C-3'; primer 4, 5'-GGC-GCG-AAG-CTT-ACT-TAA-TTG-GGC-AGG-CGC-C-3'. The entire RTPR

Table 1. Peptides of *L. leichmannii* RTPR

Peptide	Source	Sequence
1	N terminus	S <sup>2</sup> EEISLSAEFIDRVKASVKPH <sup>22</sup>
2	Active site labeling*	T <sup>112</sup> GDSLNNCWF <sup>121</sup>
3	Active site labeling*	D <sup>722</sup> LELVDQTDCEGGACPIK <sup>739</sup>
4	Peptide mapping	R <sup>449</sup> VTFSFYDWEISR <sup>461</sup>
5	Peptide mapping	V <sup>489</sup> VTGFKDDFDPEH(E) <sup>†</sup> AIKVPVYDKR <sup>513</sup>
6	Peptide mapping	S <sup>702</sup> QEITGNVEEVFSQLSDSVK <sup>721</sup>

\*Lin *et al.* (22).

<sup>†</sup>Residue in parentheses indicates possible/low confidence.

gene was amplified in two pieces using (i) primer 3 and a reverse sequencing primer 5 and (ii) primer 4 and a forward sequencing primer 6. The N-terminal half of the RTPR gene, which spanned bp 1–1275, shared a unique *Bgl* II restriction site with the C-terminal half, which spanned bp 1179–2220. Sequences of primers 5 and 6 are as follows: Primer 5, 5'-GAA-GAC-TTC-AAA-GAG-GTT-GCA-AGG-3'; primer 6, 5'-CTA-CCA-GGC-TGG-AAT-TGA-CGG-3'. Each PCR contained (in a final volume of 100  $\mu$ l) each primer at 1  $\mu$ M, 10  $\mu$ l of 10 $\times$  PCR buffer, all four dNTPs (each at 0.2 mM), and 1.2  $\mu$ g of genomic DNA. The reaction mixture was overlaid with 100  $\mu$ l of paraffin oil and heated to 94°C for 5 min. *Taq* polymerase (2.5 units) was added and 30 cycles of the following program were run: 94°C for 1 min, 45°C for 1 min, and 72°C for 3 min. The last cycle included a 10-min incubation at 72°C. Each of the two amplified fragments was ligated into the Invitrogen TA vector as described by the manufacturer. The ligation mixture was transformed into *E. coli* DH5 $\alpha$  by standard methods, and the cells were plated on SOC agar supplemented with kanamycin (50  $\mu$ g/ml) and 5-bromo-4-chloro-3-indolyl  $\beta$ -D-galactoside (0.008%). Colonies containing recombinant plasmids were identified by their white phenotype, and plasmid DNA from several white colonies was isolated and analyzed by restriction digestion and sequencing. Plasmid DNA from a colony containing the C-terminal half of the gene 3 $\alpha$ 1 was isolated and digested with *Hind*III and *Bgl* II, while plasmid DNA from a colony containing the N-terminal half of the gene 6 $\alpha$ 4 was digested with *Eco*RI and *Bgl* II. The inserts were purified by electroelution from 1.5% agarose gels, ethanol-precipitated, and desalted on Nick columns (Pharmacia).

A three-fragment ligation into pKK223-3 that had been digested with *Hind*III and *Eco*RI was carried out in a total volume of 10  $\mu$ l containing 100 ng of 3 $\alpha$ 1, 100 ng of 6 $\alpha$ 4, 200 ng of pKK223-3, 2  $\mu$ l of 5 $\times$  ligase buffer, and 1 unit of T4 DNA ligase (BRL). The reaction was incubated overnight at 16°C, and pSQUIRE, the resulting vector, was transformed into *E. coli* HB101. Constitutive expression of RTPR from pSQUIRE's *tac* promoter was observed when the bacteria were grown in LB medium. Alternatively, with pSQUIRE transformed into JM105, expression was inducible by addition of 1 mM isopropyl  $\beta$ -D-thiogalactoside to a bacterial culture at an OD<sub>600</sub> of 0.5.

**Isolation of RTPR.** A 1-ml saturated culture of *E. coli* HB101/pSQUIRE was used to inoculate 500 ml of Luria broth. The culture was grown at 37°C with aeration to late logarithmic phase (doubling time, 45 min). The bacteria were pelleted by centrifugation at 9000  $\times$  g for 25 min, and the pellet was resuspended in 20 ml of buffer A (200 mM potassium phosphate, pH 7.2/1 mM 2-mercaptoethanol/1 mM EDTA/0.1 mM phenylmethylsulfonyl fluoride) equilibrated at 4°C. All subsequent steps were performed at 4°C, and all buffers contained 1 mM 2-mercaptoethanol. The cells were disrupted by passage through a French pressure cell at 165 Pa and centrifuged again at 9000  $\times$  g for 25 min. Protamine sulfate (1% in buffer A) was added to the supernatant (22 ml) to give a final concentration of 0.25%. Addition took place over 10 min, and the resulting solution was stirred for an additional 20 min. The precipitate was removed by centrifugation at 9000  $\times$  g for 30 min. EDTA was added to the supernatant (27 ml) to a final concentration of 5 mM, and the supernatant was brought to 40% saturation with (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (243 g/liter) over 20 min. The solution was stirred an additional 50 min, and the precipitate was isolated by centrifugation at 9000  $\times$  g for 30 min. The pellet was dissolved in 1 ml of buffer B (10 mM potassium phosphate, pH 7.2/1 mM 2-mercaptoethanol) and desalted on a Sephadex G-25 (1.6  $\times$  45 cm) column, using buffer B as the eluate. The protein-containing fractions (20 ml) were pooled and loaded onto a DE52 anion-exchange column (2.5  $\times$  5.5 cm) equilibrated in

buffer B. The column was washed with 50 ml of buffer B, followed by 50 ml of 100 mM potassium phosphate (pH 7.2). RTPR was eluted with 150 mM potassium phosphate (pH 7.2). Fractions (41 ml) containing RTPR activity were pooled and exchanged into buffer C (10 mM potassium phosphate, pH 6.8/1 mM 2-mercaptoethanol) by diluting and concentrating the fractions in an Amicon ultrafiltration apparatus fitted with a PM30 membrane. The protein was loaded onto a hydroxylapatite column (2.5  $\times$  7 cm) equilibrated in buffer C, and the column was washed with 50 ml of the same buffer. The protein was eluted with a 400-ml linear gradient from 10 mM to 200 mM potassium phosphate. RTPR was eluted at 75 mM buffer. The protein fractions were pooled, concentrated to 20 mg/ml, dialyzed against 100 mM sodium citrate, pH 5.5/1 mM dithiothreitol/1 mM EDTA/0.025% sodium azide/20% (vol/vol) glycerol, and stored in aliquots at -80°C.

**Ribonucleotide Reductase Assay.** RTPR was assayed by a modification of the procedure of Blakley (23), using 54  $\mu$ M thioredoxin/64 nM thioredoxin reductase/0.14 nM RTPR.

## RESULTS

**Peptide Mapping.** RTPR was purified to near homogeneity and subjected to SDS/PAGE. The protein was electroblotted onto a poly(vinylidene difluoride) membrane or a nitrocellulose membrane. The N-terminal sequence of RTPR was determined by automated Edman degradation (Table 1) from the protein on the poly(vinylidene difluoride) membrane. Internal sequence information was obtained by digesting the nitrocellulose-bound protein with trypsin, separating the eluted peptide fragments by HPLC, and sequencing several of the fragments by automated Edman degradation (18, 19) (Table 1). Peptides 1 and 3 were used to generate the degenerate oligonucleotide primers used for the PCR. In addition, these two peptide sequences were used to design two long probes that contained inosine at the wobble base and were used for Southern blot analysis (24–26).

**Isolation of Genomic DNA and Cloning of the Gene Encoding RTPR.** Isolation of genomic DNA from *L. leichmannii* was achieved using methods developed by Chassy and Guiffrida (20), with the exception that lysozyme was used in conjunction with a second enzyme, mutanolysin. The latter was crucial for the reproducible recovery of DNA that migrated as a tight band above the 50-kb molecular size marker when subjected to electrophoresis on a 1% agarose gel.

Two methods were concurrently employed in cloning the gene for RTPR. The first consisted of constructing a size-selected *L. leichmannii* genomic DNA library and screening the library by Southern blot analysis with the oligonucleotide primers designed from our extensive peptide information (Table 1). The genomic DNA was digested with *Bam*HI, *Eco*RI, *Hind*III, or *Pst* I and screened by Southern blot analysis with the two inosine-containing probes. A *Hind*III fragment of  $\approx$ 6.6 kb was shown to hybridize under stringent conditions to both of the probes. The genomic DNA of this region was isolated by standard methods (21). However, efforts to ligate these fragments into a vector and transform them into any host were unsuccessful. Although many recombinant colonies were obtained, restriction analysis of their plasmid DNA showed no inserts of the desired size. This result, in conjunction with the appropriate controls, suggested that the native *L. leichmannii* DNA was affecting either the ligation reaction or its stability in the vector subsequent to transformation.

The second approach undertaken to clone the gene for RTPR was to use the PCR directly on the genomic DNA under nonstringent conditions. Degenerate oligonucleotide primers were designed from peptides 1 and 3. Peptide 1 corresponds to the N-terminal sequence of RTPR, and peptide 3 corresponds to a sequence that appears to be homol-

ogous to the C-terminal end of the R1 subunit of the *E. coli* reductase. If peptide 3 from the *L. leichmannii* enzyme is also located at the C-terminal end of the protein, then this PCR would conveniently result in the amplification of almost the entire gene for RTPR. The result of the electrophoretic analysis of the PCR product indicates a single band of 2.1 kb. A protein of 76 kDa [previously reported for RTPR (27)] requires a gene of  $\approx 2.07$  kb. These results suggested that a single PCR had amplified almost the entire gene for RTPR.

An attempt was made to clone the 2.1-kb PCR fragment into pUC19. Of the 18 white colonies that were screened, one had a sizable insert of 1.1 kb. Since this insert contained the information needed to initiate the sequencing process using the pUC plasmid universal primer, no further attempts were made to clone the entire PCR fragment.

**Sequencing of the Gene for RTPR.** The DNA sequence of RTPR was determined by the Sanger dideoxynucleotide sequencing method (28) using *Taq* polymerase-based double-stranded sequencing. Initial efforts focused on the pUC plasmid containing the 1.1-kb insert from the PCR. The first 60 bases read from this cloned insert encoded 20 aa (aa 8–27) of the N-terminal end of RTPR. The first 18 nt, as expected, were those that encoded the N-terminal PCR primer. More importantly, the subsequent nucleotides that were not part of the PCR primers encoded all of the amino acids observed in the N-terminal peptide sequence (Table 1).

With this exact DNA sequence information, both strands of the uncloned 2.1-kb PCR fragment were sequenced. The entire gene sequence was obtained by successively making new primers to the ends of the previously obtained sequence. The sequence of the extreme N terminus and C terminus of RTPR was determined by sequencing the 6.6-kb *HindIII* fragments of the genomic DNA. The consistency of the sequence information from these three sequencing efforts, and the consistency of the DNA sequence with peptide sequences scattered throughout the protein (Table 1), gives us a high level of confidence in the sequence.

**Expression of RTPR.** The availability of a restriction map of the RTPR gene and the recent method described by MacFerrin *et al.* (29) for maximizing translational efficiency in *E. coli* allowed formulation of the strategy for expression of RTPR. As described in detail above, the gene was cloned in two fragments from genomic DNA by using PCR. Subsequent to appropriate restriction digestion, the two fragments were ligated into digested pKK223-3 to give pSQUIRE. Upon transformation of pSQUIRE into *E. coli* HB101, expression was constitutive from the vector's *tac* promoter. Alternatively, expression could be induced by isopropyl  $\beta$ -D-thiogalactoside when pSQUIRE was transformed into *E. coli* JM105.

**Purification and Characterization of RTPR.** The procedure used to purify RTPR is very similar to that as described (9) (Table 2). From 2.5 g of *E. coli* HB101 containing pSQUIRE, 90 mg of protein (>95% pure) with a specific activity of  $1.5 \mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$  was obtained. This activity is comparable to the best reported activities of RTPR isolated from *L. leichmannii*. The apparent  $K_m$  values for ATP and adenosylcobalamin were 1.1 mM and 1.3  $\mu\text{M}$ , respectively, identical to the values determined independently from the reductase isolated from *L. leichmannii*. The recombinant protein was submitted for N-terminal sequence analysis, and its sequence was found to be identical to that predicted by the gene sequence except that the N-terminal methionine was not present. N-terminal processing was also observed in the RTPR isolated from *L. leichmannii*.

## DISCUSSION

Cloning of the gene for *L. leichmannii* RTPR resulted fortuitously from the fact that sequence information from both the

Table 2. Purification of RTPR

Step	Volume, ml	Protein, mg/ml	Specific activity, units/mg	Total units
Crude	22	30	0.35	230
Protamine sulfate	27	24	0.35	228
(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> /Sephadex G-25	20	13.6	0.63	171
DE52	41	3.2	1.0	130
Hydroxylapatite	103	0.87	1.5	134

A 500-ml culture of pSQUIRE in *E. coli* HB101 was grown to stationary phase [ $\approx 2.5$  g (wet weight)].

N-terminal and C-terminal portions of the protein was available. The C-terminal sequence is intriguing, given its putative homology to the C-terminal end of the R1 subunit of the *E. coli* RDPR. In addition, the C termini were previously implicated as being involved in catalysis by biochemical studies designed to identify the cysteines within the active site that are oxidized concomitant with substrate reduction (22). Mixed oligonucleotide probes corresponding to the C- and N-terminal ends resulted in amplification by PCR of a 2.1-kb piece of DNA that encoded almost the entire reductase gene. The few amino acids (13 at the C terminus and 7 at the N terminus) not encoded by this 2.1-kb piece were known from peptide sequence information. Given the additional peptide sequence information (Table 1) dispersed throughout the entire protein, and the fact that both strands of the DNA were sequenced, we are confident of the integrity of this sequence. This protein sequence defines a molecular mass for RTPR of 82 kDa, which contrasts with what was previously reported (76 kDa) based on SDS/gel electrophoresis and other biophysical studies (27). This value can be compared with the molecular mass of 86.5 kDa for each protomer of the R1 subunit of the *E. coli* reductase.

On the basis of extensive mechanistic studies using isotopically labeled nucleotide substrates and mechanism-based inhibitors, we have proposed that the mechanisms of the *E. coli* and *L. leichmannii* ribonucleotide reductases are very similar (1, 3). This mechanistic information in conjunction with a limited amount of protein sequence information led us to suggest that the R1 subunit of RDPR might be homologous to the monomeric RTPR (1, 22). However, a sequence homology search between these two reductases using a variety of programs, including BLAST, FASTA, and CLUSTAL, failed to reveal any statistically significant sequence alignments. A search of the entire protein data base (GenBank, Protein Identification Resource, and EMBL) failed to reveal any sequence homology between RTPR and any known protein. Furthermore, only non-*E. coli* R1 RDPRs were identified as being homologous to *E. coli* R1. A careful comparison of the protein sequence of RTPR with those of methylmalonyl-CoA mutases (30, 31) and ethanolamine ammonia lyase (32), enzymes that catalyze AdoCbl-dependent rearrangement reactions, failed to reveal any common vitamin B<sub>12</sub> binding domain. For ethanolamine ammonia lyase, some sequence similarity might have been anticipated given the similarity of its proposed mechanism (at least in the first few steps) to the postulated mechanism of RTPR (9).

While RTPR and *E. coli* RDPR do not exhibit any statistically significant sequence homology, this does not preclude the possibility that sequence fragments may be conserved by function. Our previous studies on the *E. coli* RDPR have allowed us to propose a model in which five cysteines (on the R1 subunit) are involved in nucleotide reduction. Two cysteines (positions 754 and 759) at the C-terminal end of the R1 protein are proposed to shuttle electrons from thioredoxin to the active-site cysteines (positions 225 and 462) proposed to be directly involved in nucleotide reduction (11, 12). A fifth cysteine, Cys-439, is proposed to be the protein radical

required to initiate 3' hydrogen atom abstraction from the nucleotide substrate (13). A search of the protein sequence of RTPR was, therefore, made for cysteine-containing peptides with sequence similarities to the peptides surrounding the catalytically important cysteines of the *E. coli* R1 subunit. This search revealed two fragment similarities shown below.

RTPR	DLELVQTD-C <sup>731</sup> EGGAC <sup>736</sup> PIK
RDPR	DLVPSIQDDGC <sup>754</sup> ESGAC <sup>759</sup> KI

RTPR	TNPC <sup>408</sup> GEISLA
RDPR	SNLC <sup>439</sup> LEIALP

The alignment of the cysteines in the C-terminal region of both reductases suggests that Cys-731 and Cys-736 of RTPR might function to shuttle electrons from the *in vivo* reductant (thioredoxin) to the active-site cysteines. Consistent with this proposal is the observation that the *L. leichmannii* reductase can be reduced by *E. coli* thioredoxin (22).

The alignment of Cys-408 in RTPR with a fragment containing Cys-439 of RDPR is particularly intriguing. Cys-439 in RDPR has been proposed to be the protein radical that initiates catalysis by hydrogen atom abstraction from the nucleotide substrate. Given that the putative 5'-deoxyadenosine radical generated from homolysis of the carbon-cobalt bond of AdoCbl has been shown not to directly initiate catalysis by 3' hydrogen atom abstraction (3), it is interesting to speculate that its function is to generate the thiyl radical of RTPR Cys-408, which then initiates the reduction process.

Efforts to locate by sequence gazing the two cysteines thought to be oxidized directly concomitant with substrate reduction have been unsuccessful. However, previous biochemical studies with RTPR have been interpreted to indicate that Cys-119 is close in three-dimensional space to Cys-731 and Cys-736 (22). Analogy with similar studies on the *E. coli* RDPR indicating that Cys-225 is close in three-dimensional space to Cys-754 and Cys-759 and the evidence using site-directed mutants that Cys-225 is directly involved in substrate reduction (12) suggest that Cys-119 of RTPR might be an active-site reductant. A candidate for the second cysteine of RTPR directly involved in substrate reduction, Cys-419, remains highly speculative. This cysteine is 11 aa removed from the Cys-408, while in the *E. coli* RDPR Cys-462, the second cysteine involved in substrate reduction, is 23 residues removed from Cys-439. Preliminary studies using site-directed mutants of RTPR suggest that all of these cysteine residues play important roles in catalysis (unpublished data). These results further suggest that, despite the significant differences in cofactor requirement and primary and quaternary structure, the *E. coli* and *L. leichmannii* reductases may function by surprisingly similar mechanisms.

Finally, the sequence search of RTPR revealed a third region with sequence homology to RDPR:

RTPR	Y <sup>457</sup> DWEISREIIQ
RDPR	Y <sup>599</sup> DWEALRESIK

The function of this region of RDPR is unknown, although the tryptophan is conserved in all of the iron-requiring reductases sequenced to date (6).

The cloning, sequencing, and expression of RTPR now make large amounts of this protein and the mutant proteins readily available. As RTPR is the simplest of all known reductases, detailed mechanistic and structural studies of this enzyme will provide a prototype for investigating other reductases. Crystallization of this protein is a high priority, as

is additional characterization of the wild-type protein and of the site-directed mutants.

We thank Dr. Steve Benner and Gaston Gomet of the Eidgenössische Technische Hochschule, Zurich, for help in sequence searches. This work was supported by National Institutes of Health Grant 29595 and National Institutes of Health Training Grant CA 09112 (S.B.).

- Stubbe, J. (1990) *J. Biol. Chem.* **265**, 5329–5332.
- Thelander, L. & Reichard, P. (1979) *Annu. Rev. Biochem.* **48**, 133–158.
- Stubbe, J. (1989) *Adv. Enzymol. Relat. Areas Mol. Biol.* **63**, 349–417.
- Lammers, M. & Follmann, H. (1983) *Struct. Bonding (Berlin)* **54**, 27–91.
- Blakley, R. (1965) *J. Biol. Chem.* **240**, 2173–2179.
- Eriksson, S. & Sjöberg, B. M. (1989) in *Allosteric Enzymes*, ed. Hervé, G. (CRC, Boca Raton, FL), pp. 189–215.
- Singh, D., Tamao, Y. & Blakley, R. (1977) *Adv. Enzymol. Reg.* **15**, 81.
- Nordlund, P., Sjöberg, B. M. & Eklund, H. (1990) *Nature (London)* **345**, 593–598.
- Ashley, G. W., Harris, G. & Stubbe, J. (1986) *J. Biol. Chem.* **261**, 3958–3964.
- Mao, S. S., Johnston, M. I., Bollinger, J. M. & Stubbe, J. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 1485–1490.
- Mao, S. S., Holler, T. P., Yu, G., Bollinger, J. M., Booker, S., Johnston, M. I. & Stubbe, J. (1992) *Biochemistry* **31**, 9733–9743.
- Mao, S. S., Holler, T. P., Bollinger, J. M., Yu, G., Johnston, M. I. & Stubbe, J. (1992) *Biochemistry* **31**, 9744–9751.
- Mao, S. S., Yu, G., Chalfoun, D. & Stubbe, J. (1992) *Biochemistry* **31**, 9752–9759.
- Aberg, A., Hahne, S., Karlsson, M., Larsson, A., Örmö, M., Ahgren, A. & Sjöberg, B. M. (1989) *J. Biol. Chem.* **264**, 12249–12252.
- Lunn, C. A., Kathju, S., Wallace, C., Kushner, S. & Pigiet, V. (1984) *J. Biol. Chem.* **259**, 10469–10474.
- Russel, R. & Model, P. (1985) *J. Bacteriol.* **163**, 238–242.
- Efthymiou, C. & Hansen, P. A. (1962) *J. Infect. Dis.* **110**, 258–267.
- Abersold, R. H., Leavitt, J., Saavedra, R. A., Hood, L. E. & Kent, S. B. H. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 6970–6974.
- Lane, W. S., Galat, A., Harding, M. W. & Schreiber, S. L. (1991) *J. Protein Chem.* **10**, 151–160.
- Chassy, B. M. & Guiffrida, A. (1980) *Appl. Environ. Microbiol.* **39**, 153–158.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
- Lin, A. I., Ashley, G. W. & Stubbe, J. (1987) *Biochemistry* **26**, 6905–6909.
- Blakley, R. L. (1978) *Methods Enzymol.* **51**, 246–259.
- Miura, N., Ohtsuka, E., Yamaberi, N., Ikehara, M., Uchida, T. & Okada, Y. (1985) *Gene* **38**, 271–274.
- Ohtsuka, E., Matsuki, S., Ikehara, M., Takahashi, Y. & Matsubara, K. (1985) *J. Biol. Chem.* **260**, 2605–2608.
- Takahashi, Y., Kato, K., Hayashizaki, Y., Wakabayashi, T., Ohtsuka, E., Matsuki, S., Ikehara, M. & Matsubara, K. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 1931–1935.
- Panagou, D., Orr, M. D., Dunstone, J. R. & Blakley, R. L. (1972) *Biochemistry* **11**, 2378–2388.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
- MacFerrin, K. D., Terranova, M. P., Schreiber, S. L. & Verdine, G. L. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 1937–1941.
- Ledley, F. D., Lumetta, M., Nguyen, P. N., Kolhouse, J. F. & Allen, R. H. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 3518–3526.
- Marsh, E. N., McKie, N., Davis, N. K. & Leadlay, P. F. (1989) *Biochem. J.* **260**, 345–352.
- Faust, L. P., Connor, J. A., Roof, D. M., Hoch, J. A. & Babior, B. M. (1990) *J. Biol. Chem.* **265**, 12462–12466.